Understanding Type 1 and Type 2 Errors from the Feline Perspective: All Mistakes Are Not Equal!

Stats Cat · 25 January, 2013

Serving cat food? I sure hope you've set your alpha level high enough.

"Bad kitty!" That's a phrase you almost never hear, but even we cats make the occasional mistake. I was reminded of this recently as I watched my human trying to analyze some data. People frequently make mistakes when they test a hypothesis with data analysis. Specifically, they can make either Type I or Type II errors.

When I first started reading my human's statistics textbooks a few years ago, this idea seemed awfully silly to me. We cats appreciate being direct, and you either get the answer correct or you don't. I mean, a mistake is a mistake, right? Does it really matter *how* you reached the wrong conclusion?

Then I recalled how, as a kitten, I once made a mistake near my litter box. Later that week, I made a mistake on my human's bed. While these two mistakes had certain similarities, the responses they elicited from my human couldn't have been more different. So even as a kitten, I learned that not all mistakes have the same impact.

If you're talking about statistical questions, the way you make an error certainly *does* matter.

Type I and Type II Errors: A Vital Example of Why It Matters

Textbook authors throw around lots of examples about Type I and Type II errors and why they're important. They'll cite allegedly life-threatening examples that involve, say, testing the effectiveness of different medicines. Or the reliability of airplane parts. Or the stopping distances for different brands of car tires.

Whatever. I guess humans care about that sort of thing, but let's be honest--none of those examples are worth a mummified mouse tail to any self-respecting cat.

So let's talk about Type I and Type II errors as they apply to a situation that actually matters, one where lives *really* hang in the balance: the taste of cat food.

Assume that the Puma Gourmet cat food company wants to compare two formulations for a new food. The null and alternative hypotheses are:

Null hypothesis (Ho): m1 = m2, or "Both types of cat food taste the same."

Alternative hypothesis (H1): m1 ≠ m2, or "Both types of cat food do not taste the same."

The company's brainiest science guys are assigned to conduct some taste tests, gather data about how well a representative sample of cats likes each formulation, and then analyze the data.

If the conclusion they reach matches the reality of the situation, wonderful. Can't wait to try the new food. But what if they've made a mistake, even inadvertently?

To Reject, or Fail to Reject, The Null Hypothesis

As in any hypothesis test, the Puma Gourmet researchers must decide whether or not to reject the null hypothesis based on the data they've collected.

So when we talk about Type I and Type II errors, we're really talking about the two different ways in which you can botch the decision whether or not to reject the assumption that the null hypothesis (Ho) is correct. Rejecting the null hypothesis when it is true is a Type I error. Failing to reject the null hypothesis when it is false is a Type II error.

It's a little easier to understand if you look at it in tabular form:

The Reality	*We do not reject the null (Ho)*	*We reject the null (Ho)*
Null (Ho) is true.	Correct decision. Well done!	Type I error
Null (Ho) is false.	Type II error	Correct decision. Sweet!

If you make a Type I error, you reject the null hypothesis when, in fact, it's true. In our example, the company would conclude that the two cat food formulations taste different when they really don't. Since the cat foods taste the same, this error is not a complete disaster, because at least the cats will experience the same great Puma Gourmet taste regardless of which formulation they get.

If you commit a Type II error, you do not reject the null hypothesis even though it is false. In this case, the company would conclude that the cat foods taste the same when, in fact, they taste different. Imagine how devastating this error would be if, as a result, cats got a less delicious version of their Puma Gourmet food.

Choosing the Right Amount of Risk for Your Statistical Analysis

Now you can see why it's important to understand the difference between Type I and Type II errors as you conduct your own hypothesis tests. For instance, if you're working on a Six Sigma project that could save your company millions, is the probability of committing one type of error more serious or costly than committing the other type?

This ties in to the statistical concepts of risk, significance, and power. Statisticians refer to the probability of making a Type I error as "alpha," or the"significance level" you set for your hypothesis test. A common default value for alpha is 0.05, which means you have a 5 percent chance of rejecting the null hypothesis when it is true. A lower alpha value gives you a lower risk of incorrectly rejecting the null hypothesis. When it's really important, like in our cat food example, researchers will select an alpha value of 0.01, which would reduce the chances of a Type I error to just 1 percent.

The concept of "power" is related to the probability of making a Type II error. Statisticians refer to this as "beta," and it's a value that statisticians typically cannot know. But researchers can lessen the risk of Type II errors by making sure their tests have enough power -- in other words, by making sure their sample size is large enough to detect a difference when one truly exists.

I hope this explanation has helped you appreciate the importance of avoiding both types of errors. And if you're one of those Puma Gourmet taste researchers, I hope you'll take this to heart!

You Might Also Like:

Understanding Hypothesis Tests: Significance Levels (Alpha) and P values in Statistics
Bewildering Things Statisticians Say: "Failure to Reject the Null Hypothesis"
Why Shrewd Experts "Fail to Reject the Null" Every Time
Alphas, P-Values, and Confidence Intervals, Oh My!

Comments

Name: Jim Taylor, CRE, CPE, CPMM • Friday, January 25, 2013

One of the best explanations I've seen. Now explain how to find out how big the sample is to have enough power.

Thanks,

Jim

Name: Carmen Frost • Wednesday, January 30, 2013

Dear Stat Cat,

We all make mistakes... Thank goodness you know how to set your alpha levels high enough!

I hope in the near future you can write a stat article about mice. My favourite topic!

Have you ever had MICE-TEA?

Name: Cathy Edwards • Wednesday, February 12, 2014

That was awesome! I feel so much smarter! ;)